

SYMPLECTIC AND LIE ALGEBRAIC TECHNIQUES IN GEOMETRIC OPTICS

J.F. Cariñena[†], C. López[‡] and J. Nasarre^{*}

[†]Departamento de Física Teórica, Universidad de Zaragoza, Zaragoza 50.009

[‡] Departamento de Matemática Aplicada, CPSI, Universidad de Zaragoza, Zaragoza 50.015

^{*}IES Miguel Catalán, Isabel la Católica 3, Zaragoza 50.009

Abstract

We will show the usefulness of the tools of Symplectic and Presymplectic Geometry and the corresponding Lie algebraic methods in different problems in Geometric Optics.

1 Introduction: Symplectic and Presymplectic geometry

Geometric techniques have been applied to physics for more than 50 years in many different ways and they have provided powerful methods of dealing with classical problems from a new geometric perspective. Linear representations of groups, vector fields, forms, exterior differential calculus, Lie groups, fibre bundles, connections and Riemannian Geometry, symmetry and reduction of differential equations, etc..., are now well established tools in modern physics. Now, after more than twenty years of using Lie algebraic methods in Optics by Dragt, Forest, Sternberg, Wolf and their coworkers, we aim here to establish the appropriate geometric setting for Geometric Optics. Applications in computation of aberrations for different orders will also be pointed out.

The basic geometric structure for the description of classical (and even quantum) systems is that of symplectic manifold. A *symplectic manifold* is a pair (M, ω) where ω is a nondegenerated closed 2-form in M . If ω is exact we will say that (M, ω) is an *exact symplectic manifold*. Let $\hat{\omega} : \mathfrak{X}(M) \rightarrow \Lambda^1(M)$ be given by $\hat{\omega}(X) = i(X)\omega, \hat{\omega}(X)Y = \omega(X, Y)$. The two-form ω is said to be nondegenerate when $\hat{\omega}$ is a bijective map. Then M is evendimensional and it may be used to identify vector fields on M with 1-forms on M . Vector fields X_H corresponding to exact 1-forms dH are called Hamiltonian vector fields. The 2-form ω is said to be closed if $d\omega = 0$.

The simplest example is \mathbb{R}^{2n} with coordinates $(q^1, \dots, q^n, p_1, \dots, p_n)$ endowed with the constant 2-form $\omega = \sum_{i=1}^n dq^i \wedge dp_i$. Closedness of ω is very important because Darboux theorem establishes that for any point $u \in M$ there exists a local chart (U, ϕ) such that if $\phi = (q^1, \dots, q^n; p_1, \dots, p_n)$, then $\omega|_U = \sum_{i=1}^n dq^i \wedge dp_i$. Consequently, the example above is the local prototype of a symplectic manifold. It is also well known that if Q is the configuration space of a system, its cotangent bundle, $T^*Q = \bigcup_{q \in Q} T_q^*Q$, called phase space, is endowed with a canonical 1-form θ on T^*Q such that $(T^*Q, -d\theta)$ is an exact symplectic manifold. More especcially, if (q^1, \dots, q^n) are coordinates in Q then $(q^1, \dots, q^n, p_1, \dots, p_n)$ are coordinates in T^*Q and $\theta = \sum_{i=1}^n p_i dq^i, \omega = \sum_{i=1}^n dq^i \wedge dp_i$.

A Hamiltonian dynamical systems is a triplet (M, ω, H) where M is a differentiable manifold, $\omega \in Z^2(M)$ is a symplectic form in M and $H \in C^\infty(M)$ is a function called Hamiltonian. The

dynamical vector field X_H is then the solution of the equation $i(X_H)\omega = dH$. In the example above

$$X_H = \frac{\partial H}{\partial p_i} \frac{\partial}{\partial q^i} - \frac{\partial H}{\partial q^i} \frac{\partial}{\partial p_i}.$$

and the same expression is valid in Darboux coordinates for any Hamiltonian dynamical system. A Poisson bracket can be introduced in a symplectic manifold (M, ω) by $\{F, G\} = X_G F = \omega(X_F, X_G)$. Then, closedness of ω is equivalent to Jacobi identity for P.B. Moreover, it may be shown that $\sigma = \hat{\omega}^{-1} \circ d : C^\infty(M) \rightarrow \mathfrak{X}(M)$ is a Lie algebra homomorphism.

A presymplectic manifold is a pair (M, ω) such that M is a differentiable manifold and ω is a constant rank closed 2-form in M . The kernel of ω defines an integrable distribution (because $d\omega = 0$) and when the set of leaves is a manifold, it can be endowed with a symplectic structure. This process is called reduction of the presymplectic structure.

Very interesting examples of HDS are those defined by regular Lagrangians, (TQ, ω_L, E_L) , with $\omega_L = -d\theta_L = -d(dL \circ S)$, $E_L = \Delta L - L$. More accurately, the geometric approach to the Lagrangian description makes use of the geometry of the tangent bundle of the configuration space that we will shortly review. The tangent bundle $\tau_Q : TQ \rightarrow Q$ is characterized by the existence of a vector field generating dilations along the fibres, called Liouville vector field, $\Delta \in \mathfrak{X}(TQ)$, and the vertical endomorphism which is a $(1, 1)$ -tensor field S in TQ that in a natural coordinate system for TQ , induced from a chart in Q , are $\Delta = v^i \partial/\partial v^i$, and $S = (\partial/\partial v^i) \otimes dq^i$. Given a function $L \in C^\infty(TQ)$, we define the 1-form $\theta_L \in \Lambda^1(TQ)$ by $\theta_L = dL \circ S$. When the exact 2-form $\omega_L = -d\theta_L$ is nondegenerate the Lagrangain L is called regular and then (TQ, ω_L) is a symplectic manifold. The energy function E_L is given by $E_L = \Delta(L) - L$. The coordinate expressions are $\theta_L = (\partial L/\partial v^i)dq^i$ and $E_L = v^i(\partial L/\partial v^i) - L$.

2 Symplectic structures in Geometric Optics

The set of oriented geodesics of a Riemannian manifold can be endowed with a symplectic structure and in particular the set of oriented straightlines in the plane, which is the set of light rays in a two-dimensional constant rank medium, can be endowed with a symplectic structure. Moreover, it can be considered as the cotangent bundle of the one-dimensional sphere S^1 . If an origin O has been chosen in the plane, every oriented straightline that does not pass through the point O is characterized by a unit vector \mathbf{s} pointing in the line direction and a vector \mathbf{v} orthogonal to \mathbf{s} with end on the line and origin in O . The straightlines of a pencil of oriented parallel lines are characterized by proportional vectors \mathbf{v} and the same \mathbf{s} . Straightlines passing through O with direction given by \mathbf{s} correspond to $\mathbf{v} = \mathbf{0}$. The vectors \mathbf{v} and \mathbf{s} being orthogonal and $\mathbf{s} \cdot \mathbf{s} = 1$, the couple (\mathbf{s}, \mathbf{v}) can be seen as a tangent vector to the unit circle S^1 at the point described by \mathbf{s} .

The Riemannian metric in S^1 can be used to identify in each point \mathbf{s} the tangent space $T_{\mathbf{s}}S^1$ with its dual space $T_{\mathbf{s}}^*S^1$ and therefore the tangent bundle TS^1 with the cotangent bundle T^*S^1 . This identification shows us that the space of oriented straightlines in the Euclidean two-dimensional space can be endowed with an exact symplectic structure which corresponds to the canonical structure for the cotangent bundle T^*S^1 . The study of oriented straightlines in Euclidean three-dimensional space follows a similar pattern.

A choice of coordinates in the base space will provide us Darboux coordinates: a good choice will be an angle coordinate. A straightline $y = m x + b$ with slope $m = \tan \theta$ will be represented by

a vector orthogonal to the vector $\mathbf{s} = (\cos \theta, \sin \theta)$, and length $b \cos \theta$, namely, $\mathbf{v} = b \cos \theta \frac{\partial}{\partial \theta}$. The vector $\partial/\partial \theta$ is unitary in the Euclidean metric, and then the point $(\theta, p_\theta) \in T^*S^1$ corresponding to (θ, v_θ) is $p_\theta = v_\theta$. The symplectic form in TS^1 translated from the canonical symplectic structure in T^*S^1 $\omega_0 = d\theta \wedge dp_\theta$ will be $\omega = d\theta \wedge d(b \cos \theta) = d(\sin \theta) \wedge db$. Therefore, Darboux coordinates for ω adapted to the cotangent structure are not only $(\theta, b \cos \theta)$ but also $q = \sin \theta$, $p = b$, which are more appropriate from the experimental viewpoint. So, the flat screens arise here as a good choice for Darboux coordinates.

The choice usually done in Geometric Optics is $\mathbf{s} \cdot \mathbf{s} = n^2$, the Darboux coordinate q then being $q = n \sin \theta$. This leads to the image of the Descartes sphere, a sphere of radius n whose points describe the ray directions. In the more general case of a variable refractive index, we recall that light rays trajectories in Geometric Optics are determined by Fermat's principle: the ray path connecting two points is the one making stationary the optical length: $\delta \int_\gamma n ds = 0$.

This corresponds to the well-known Hamilton's principle of Classical Mechanics with an “optical Lagrangian” $L = n \sqrt{v_x^2 + v_y^2 + v_z^2}$, which is a differentiable function in $T\mathbb{R}^3$ up to the zero section. In other words, the mechanical problem corresponding to Fermat's principle leads to a singular Lagrangian $L(q, v) = [g(v, v)]^{1/2}$, where g is a metric conformal to the Euclidean metric g_0 , $g(v, w) = n^2 g_0(v, w)$. L is an homogeneous function of degree one in the velocities and consequently L is singular and the corresponding energy function vanishes identically.

It was shown in [1] that it is possible to relate the solutions of the Euler–Lagrange equations for L with those of the regular Lagrangian $\mathbb{L} = \frac{1}{2}L^2$, up to a reparametrization. \mathbb{L} is quadratic in velocities and the solution $\Gamma_{\mathbb{L}}$ of the equation $i(\Gamma_{\mathbb{L}})\omega_{\mathbb{L}} = dE_{\mathbb{L}} = d\mathbb{L}$ is not only a second order differential equation vector field but also a spray, the projection onto \mathbb{R}^3 of its integral curves being the geodesics of the Levi–Civita connection defined by g . The kernel of ω_L is two-dimensional and it is generated by $\Gamma_{\mathbb{L}}$ and the Liouville vector field Δ . The distribution $\ker \omega_L$ is integrable because $d\omega_L = 0$; the distribution is also generated by Δ and $K = \frac{1}{v^3}\Gamma_{\mathbb{L}}$, for which $[\Delta, K] = 0$.

If the refractive index for an optical system depends only on x^3 and the region in which the index is not constant is bounded, we can choose Darboux coordinates by fixing a x^3 outside this region and taking Darboux coordinates for the corresponding problem of constant index [2]. This justify the choice of coordinates for the ingoing and outgoing light rays in the constant index media, i.e. it shows the convenience of using flat screens in far enough regions on the left and right respectively, and then this change of Darboux coordinates seems to be, from an active viewpoint, a canonical transformation. Similar results can be obtained (see [3]) for nonisotropic media for which the refractive index depends only on the ray direction, i.e. $n = n(v)$ and $\Delta n = 0$. The only difference is that $\omega_{\mathbb{L}}$ may be singular, but in the regular case all works properly.

3 Group theoretical approximations

Mathematical expressions like $x' = f(x)$ admit two different interpretations. In the *alias* interpretation x and x' are coordinates of the same point in two different coordinate systems, while in the *ad libi* interpretation x are the coordinates of a point and x' those of its image under the transformation defined by f . In this sense a change of Darboux coordinates can be seen as a canonical transformation in \mathbb{R}^{2n} , and in particular, when Darboux coordinates are chosen as indicated above, the passage of the set of light-rays through an optical device can be considered

as a canonical transformation. Moreover, we can split an optical system in two subsystems and the canonical transformation factorizes as a product of two canonical transformations. Even if the group of canonical transformations is not a Lie group, any element g can be written as the exponential of an element in its Lie algebra, the set of Hamiltonian systems, $g = \exp X_f$. Symmetry of the optical system leads to reduction, and then to a lower number of degrees of freedom.

There exist formulae generalizing Baker-Campbell-Hausdorff for composition of generating functions, both in an exact or approximate way. Most of approximation formulae substitute the generating functions by a power series development and then only keep some terms, giving rise in this way to aberrations. For instance, if we only consider quadratic terms, we will get the linear approximation.

The fundamental algebraic ingredients for the theory of approximate groups are the concepts of enveloping algebra \mathcal{U} and symmetric algebra Σ of a Lie algebra \mathfrak{g} . Essentially, if $\{X_1, \dots, X_n\}$ is a basis of \mathfrak{g} , then Σ is the algebra of polynomials in $\{X_1, \dots, X_n\}$. Both \mathcal{U} and Σ have graded Lie algebra structures extending that of \mathfrak{g} , which can be identified as a subalgebra of \mathcal{U} and Σ . In the same way as \mathfrak{g} can be seen as the set of linear functions on \mathfrak{g}^* , the symmetric algebra can be considered as the set of polynomials on \mathfrak{g}^* . The adjoint action of G on \mathfrak{g} can be extended to an action $\text{Ad} : G \times \Sigma \rightarrow \Sigma$ in such way that $\text{Ad}(g)$ is linear for each $g \in G$ and $\text{Ad}(g)(p_1 \cdot p_2) = \text{Ad}(g)(p_1) \cdot \text{Ad}(g)(p_2)$. The extension of the adjoint action of the Lie algebra, $\text{ad} : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$, $\text{ad}(a, b) = \text{ad}(a)(b) = [a, b]$, is the adjoint action of the symmetric algebra Σ :

$$\text{ad} : \Sigma \times \Sigma \rightarrow \Sigma, \quad \text{ad}(p_1, p_2) = \text{ad}(p_1)(p_2) = [p_1, p_2]_{\Sigma}.$$

For any $p \in \Sigma$, we can also consider the formal transformation of Σ , $\phi(p) : \Sigma \rightarrow \Sigma$,

$$p' \rightarrow \phi(p)(p') = \exp(\text{ad}(p))(p') = p' + [p, p']_{\Sigma} + 1/2[p, [p, p']_{\Sigma}]_{\Sigma} + \dots,$$

and we should now consider the elements p for which such expression is meaningful. They span a group G_{Σ} . The enlarged action of it reduces to the identity when acting on the set $\Sigma^I = \{p \in \Sigma \mid [a, p]_{\Sigma} = 0, \forall a \in \mathfrak{g}\}$ of the polynomial Casimir elements of \mathfrak{g} . We shall then pass to the quotient graded Lie algebra $\Sigma^C = \Sigma/\Sigma^I$. Finally for approximation we will consider for each $r \in \mathbb{N}$ the ideal spanned by $H_r = \bigoplus_{t > r} \Sigma_t$, and then $\phi([p])H_r \subset H_r$, and therefore it induces a map $\Phi^r([p])$.

4 Perturbative treatment of aberrating optical systems using Weyl group

A model for geometrical optics in a plane is obtained from the Weyl group $W(1)$: it is a three-dimensional Lie group, with elements $g \in W(1)$ labelled by $g = (\mu_1; \nu)$, $\mu_1 \in \mathbb{R}^2$, $\nu \in \mathbb{R}$, and composition law $g'g = (\mu_1 + \mu'_1, \nu' + \nu + \frac{1}{2}\mu'_1 \wedge \mu_1)$ where \wedge denotes $(a, b) \wedge (c, d) = ad - bc$.

A basis for the Lie algebra $\mathfrak{w}(1)$, is given by $Q = \partial_a - \frac{1}{2}b\partial_{\nu}$, $P = \partial_b - \frac{1}{2}a\partial_{\nu}$, $I = \partial_{\nu}$ with Lie brackets $[Q, I] = [P, I] = 0$, $[Q, P] = I$. An infinite-dimensional basis for the associated symmetric algebra Σ is given by $\{1, I, Q, P, I^2, IQ, IP, Q^2, QP, P^2, \dots\}$. $\phi(\lambda_1 I + \lambda_2 Q + \lambda_3 P)$ is in fact an element of G . Its generalized adjoint action on Σ preserves each subspace Σ_r . It is enough to know its action on Σ_1 , $(c_1 I + c_2 Q + c_3 P \rightarrow ((c_1 - \lambda_3 c_2 + \lambda_2 c_3)I + c_2 Q + c_3 P))$. Another typical element of G_{Σ} is given by $\phi(\nu_1 I^2 + \nu_2 IQ + \nu_3 IP + \nu_4 Q^2 + \nu_5 QP + \nu_6 P^2)$, which in fact is only a

formal map. However, its projected maps on \mathbb{P}_r are well defined. The infinitesimal adjoint action $\text{ad}(\nu_1 I^2 + \nu_2 IQ + \nu_3 IP + \nu_4 Q^2 + \nu_5 QP + \nu_6 P^2)$ maps each Σ_r onto Σ_{r+1} , so that only its action on Σ_1 is not trivial when considering its projected map on $\mathbb{P}_2 = \Sigma_0 \oplus \Sigma_1 \oplus \Sigma_2$.

The Casimir elements of $\mathfrak{w}(1)$ are the polynomial functions on I and 1. A basis for the quotient algebra Σ^C is $\{[1]^c, [Q]^c, [P]^c, [Q^2]^c, [QP]^c, [P^2]^c, \dots\}$. The reduction process can be obtained by quotient by the ideal generated by the Casimir I . A typical element of the group G_{Σ^C} is $\phi^C(\nu_1 [Q^2]^c + \nu_2 [QP]^c + \nu_3 [P^2]^c)$. Its projected action on \mathbb{P}_2^C is given by the matrix

$$\phi^C(\nu_1 [Q^2]^c + \nu_2 [QP]^c + \nu_3 [P^2]^c) = \begin{pmatrix} 1 & 0_{1 \times 2} & 0_{1 \times 3} \\ 0_{2 \times 1} & M & 0_{2 \times 3} \\ 0_{3 \times 1} & 0_{3 \times 2} & D^2(M) \end{pmatrix}$$

where $M \in SL(2, \mathbb{R})$ and $D^2(M)$ is the image of M in the three-dimensional representation:

$$M = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}, \quad D^2(M) = \begin{pmatrix} \alpha^2 & \alpha\beta & \beta^2 \\ 2\alpha\gamma & \alpha\delta + \beta\gamma & 2\beta\delta \\ \gamma^2 & \gamma\delta & \delta^2 \end{pmatrix}.$$

The matrix M associated to the map $\phi^C(\nu_1 [Q^2]^c + \nu_2 [QP]^c + \nu_3 [P^2]^c)$ is given by

$$M = \begin{pmatrix} \cosh \omega - \frac{\nu_2}{\omega} \sinh \omega & 2\frac{\nu_1}{\omega} \sinh \omega \\ -2\frac{\nu_3}{\omega} \sinh \omega & \cosh \omega + \frac{\nu_2}{\omega} \sinh \omega \end{pmatrix}, \text{ with } \omega = \pm \sqrt{\nu_2^2 - 4\nu_1\nu_3}.$$

A matrix representation for the group $G_{\Sigma_r^c}$ provides a perturbative treatment of $(r-1)$ -th order (in 1+1 dimensions). Using the factorization theorem and the axial symmetry of the system, only transformations of type $\phi^C([p]^c)$ with $[p]^c \in \Sigma_2^C$ or Σ_4^C must be considered for third order aberrations. $\phi^C(\nu_1 [Q^2]^c + \nu_2 [QP]^c + \nu_3 [P^2]^c)$ has a matrix representation immediate generalization of the above mentioned on \mathbb{P}_2^C . The element $\phi^4(\mu_1 [Q^4]^c + \mu_2 [Q^3 P]^c + \dots + \mu_5 [P^4]^c)$ is

$$\phi^4(\mu_1 [Q^4]^c + \mu_2 [Q^3 P]^c + \dots + \mu_5 [P^4]^c) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 \\ 0 & M_1 & 0 & I & 0 \\ 0 & 0 & M_2 & 0 & I \end{pmatrix}.$$

$$M_1 = \begin{pmatrix} -\mu_2 & 4\mu_1 \\ -2\mu_3 & 3\mu_2 \\ -3\mu_4 & 2\mu_3 \\ -4\mu_5 & \mu_4 \end{pmatrix}, \quad M_2 = \begin{pmatrix} -2\mu_2 & 4\mu_1 & 0 \\ -4\mu_3 & 2\mu_2 & 8\mu_1 \\ -6\mu_4 & 0 & 6\mu_2 \\ -8\mu_5 & -2\mu_4 & 4\mu_3 \\ 0 & -4\mu_5 & 2\mu_4 \end{pmatrix}$$

The representation splits into two, acting respectively on the even and odd degree subspaces. The first one can be used to find the composition law and the second one can be used to find the approximate coadjoint action on the coordinate functions q and p .

The free propagation, till the third degree approximation is $p' = p$, $q' = q + \frac{z}{n}p + \frac{z}{2n^3}p^3$, and the group element will be of the form $\phi([p_2]^c) \circ \phi([p_4]^c)$, with $[p_i]^c \in \Sigma_i^c$. The matrix $M \in SL(2, \mathbb{R})$ associated to $\phi([p_2]^c)$ is $M = \begin{pmatrix} 1 & 0 \\ z/n & 1 \end{pmatrix}$ while $[p_4]^c = -\frac{z}{8n^3}[P^4]^c$, so that, $\{\mu_1, \mu_2, \mu_3, \mu_4, \mu_5\} = \{0, 0, 0, 0, -\frac{z}{8n^3}\}$ determines M_1 and M_2 .

The same calculus for a refracting surface gives

$$M = \begin{pmatrix} 1 & \frac{n_1 - n_2}{R} \\ 0 & 1 \end{pmatrix}, \text{ and } \mu = \left\{ \frac{3n_2 - n_1 - 2\frac{n_2^2}{n_1}}{8R^3}, \frac{1 - \frac{n_2}{n_1}}{2R^2}, \frac{\frac{1}{n_2} - \frac{1}{n_1}}{4R}, 0, 0 \right\}.$$

5 Example

In the design of a doublet we have seven of these basic systems concatenated so that the total system (chosen to be telescopic) is obtained in third order approximation by composition formulae of the corresponding third order aberration group. The composition of the systems (M_2, μ_2) and (M_1, μ_1) , with M_i the linear approximation matrices and μ_i the coefficients of the fourth order polynomials, is obtained by the formula $(M_2 M_1, D^4(M_1^{-1})\mu_2 + \mu_1)$, $D^4(M)$ being the former representation of matrix M on the fourth order polynomial space. In our example, a concatenation of compositions for the doublet gives way to a total linear approximation matrix (on which we can impose the telescopic condition $\gamma = 0$ and a given factor of magnification, say $\delta = 5$) and a total fourth order polynomial. Fixing the refractive indexes of the lenses as $n_1 = 7/4$ and $n_2 = 9/4$ we obtain a seven parameter system of equations. The polynomic expression of μ in terms of S_i and z_j (the radii and lengths of the lenses) is a set of eleven to sixteen degree polynomials, so that numerical calculus should be used to obtain solutions with zero third order aberrations.

The example does not try to be a realistic design, in which a four dimensional space should be used on the q 's and p 's, chromatic aberration should be taken into account through the dependence on the refractive index with the wave length, and stability of the solution to errors on the parameters should be considered. For our playing design system, and in order to simplify the calculus, we can fix some of the parameters in terms of the other ones, say $S_2 = S_3 = z_1/4$, $z_3 = z_1$ and $z_2 = 2z_1$, so that only three parameters are left free. Taking into account the telescopic condition and the given factor of magnification we are left with just one parameter, which can be used to minimize the square of the μ vector.

A numerical solution obtained for this simplified case is $z_1 = 0.5925$, which gives a linear approximations for the total system and a total vector μ_{total}

$$M_{\text{total}} = \begin{pmatrix} 0.2 & 1.7587 \\ 2.98 \cdot 10^{-19} & 5. \end{pmatrix}, \mu_{\text{total}} = \{0.1295, -1.4124, 1.2882, 0.7881, -2.3739\}.$$

Acknowledgments

JFC and CLL acknowledge partial financial support from DGICYT under project PB-93.0582

References

- [1] J. F. Cariñena and C. López, Int. J. Mod. Phys. **6**, 431 (1991).
- [2] J. F. Cariñena and J. Nasarre, Forts. der Phys. **44**, 181 (1996).
- [3] J. F. Cariñena and J. Nasarre, J. Phys. A: Math. Gen. **29**, 1695 (1996).